Computerized processing of editorial copy

H. D. Parks

Time Inc. is the publisher of three weeklies, *Time, Life,* and *Sports Illustrated,* with a combined circulation of about 12 million copies, a monthly magazine, *Fortune,* and a considerable number of books, mainly non-fiction. Almost all of the editorial content of these publications is produced by the company's staff of writers and editors. On the other hand, almost all of the company's production work is done by contract printers.

Copy is very carefully fitted before it is sent to the printer, and as a consequence justified perforated tape is sent to printers Since teletypesetting was developed specifically for use with Linotype, *Time,* which uses Linotype almost exclusively, can be set from the reperforation of the same tape used to transmit copy to the printer. On the other hand, the primary typesetting method used by *Life* and *Sports Illustrated* is Monotype and prior to the advent of the computer there has been no economical way of converting the six level tape into 31 level ribbon for the casters. Therefore, the entire job of keyboarding had to be repeated at the printer. Finally, Time Inc. has its own photo typesetting operation using three Photon 200 equipped to take TTS tape, and this method of composition is used almost entirely by the book division.

Although computed justification and hyphenation will be contributory, the real value of the new system will come from its ability to store in its memory copy which will then be available for later use. This is the so-called editing function. It is characteristic of magazine and book editorial production systems, particularly where the material originates on the publisher's premises, that each story or caption or other unit of text matter is edited and re-edited through many cycles, very often with complete retyping of the copy between successive editing functions. It is not unusual for *Time* copy to be completely retranscribed four times before it is final. This is 'author's alterations' with a vengeance. A consequence of this rewriting and editing activity is that often the major work remains to be done when there is the least amount of time left to do it. In this environment a system which is capable of readily combining the sound portions of earlier versions with the latest corrections and producing justified and hyphenated copy can be of considerable value.

The editing-processing cycle

Fig 1 illustrates the editorial production cycle using a computer. The process begins with the writer preparing manuscript copy on his own machine. This copy is sent to a keyboard operator who completely transcribes the manuscript. Currently this is done on typewriters producing multiple forms which are returned for editing. In the new system, as the illustration shows, the useful product of the keyboard is the coded signals which are fed to the copy processor. During the first cycle 100 per cent of the information comes from the keyboard, and these characters are combined into justified lines and printed out on the high-speed printer. The output of the high-speed printer will appear to the editor to be much the same as that now obtained from typewriters including the use of upper and lower case. The main difference will be the presence of line numbers generated by the computer and placed along the left side margin.

The copy is sent to the editors and in the course of time the edited version is returned to the copy processor facility. Now the keyboard operator has the opportunity of employing the lines stored in the copy memory unit in producing the new version of the story. This is the basic cycle of the system and for each cycle the computer recalculates the justification after merging the characters from new lines with those from saved lines. Each new version is stored in the copy memory along with all the earlier versions which are still available should the editor wish to return to one of them. Eventually the decision will be made to send the latest version to the printer and, with or without further corrections, a keyboard message will cause the copy processor to drive the tape punch to produce a typesetting tape.

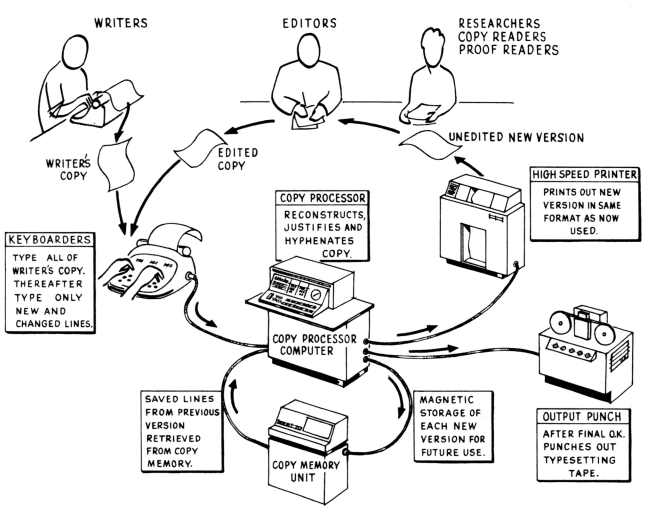
WRITERS
EDITORS
RESEARCHERS
COPY READERS
PROOF READERS
WRITER'S COPY
EDITED COPY
UNEDITED NEW VERSION
HIGH SPEED PRINTER
PRINTS OUT NEW VERSION IN SAME FORMAT AS NOW USED.
COPY PROCESSOR
RECONSTRUCTS, JUSTIFIES AND HYPHENATES COPY.
KEYBOARDERS
TYPE ALL OF WRITER'S COPY. THEREAFTER TYPE ONLY NEW AND CHANGED LINES.
COPY PROCESSOR COMPUTER
SAVED LINES FROM PREVIOUS VERSION RETRIEVED FROM COPY MEMORY.
COPY MEMORY UNIT
MAGNETIC STORAGE OF EACH NEW VERSION FOR FUTURE USE.
OUTPUT PUNCH
AFTER FINAL O.K. PUNCHES OUT TYPESETTING TAPE.

Fig 1 Editorial production cycle using copy processor system

Output

The computer programs are divided into major units which are oriented to the kind of output desired. The nature of the desired

output is indicated on a copy sent to the input keyboard and the operator informs the computer of the nature of the output and other typographic parameters by virtue of the first line keyboarded. This line is known as the index line. Five kinds of output are being programmed for. These are editorial and production hard copy, teletypesetter tape, Monotype conversion tape, and a special Photon tape. The high-speed printer will have a set of 120 different characters to be called out by the program. This, of course, permits an even more expressive hard copy than can be obtained with a commercial typewriter and full advantage will be taken of this fact. Eventually the program will be able to call for diacritical marks as well as other special characters. This form of output does have one disadvantage. It is more meaningful than the teletypesetting tapes used for transmission to the printers. To be certain that proofreaders do not assume more information being sent to the printer than is physically possible, proofreading and dummying will be done with the production hard copy. Although this will come off the same high-speed printer as the editorial hard copy, the character set used in production hard copy will be limited to those which can be specified directly from the codes in a TTS tape.

Since the inception of the project there has been a plan to convert to automatic operation of casting machines beginning with the development of a special six level code which could be easily converted to the conventional 31 level ribbon to drive a Monotype caster. The Monotype conversion unit has removed the need for the development of such a device.

The remaining kind of output is Photon Direct. At present the 200s are equipped to read teletypesetter tape in lieu of direct operation of the Photon keyboard. There is much to be gained from bypassing the keyboard and the electro-mechanical memories and computing circuits used by the front end of the Photon 200. The design of the photo typesetter is such that it uses internally a 15 level code which is most explicit in the sense that it identifies virtually every character, character width, and machine function which the machine is capable of performing. The application proposed here is quite similar to that used by Michael Barnett in working with the Photon 560 which is essentially the business end of the 200. To provide this code, the computer will punch out three frames of six level information for each character set.

General processing scheme

Fig 2 shows the flow of information through the computer, into and out of storage, and finally to the high-speed printer or punch. Characters coming from the input tape readers will be stored in sequence in the input character area. As each character in this area is examined in turn it is used to address an entry in the main character table. This table serves a dual role. In the case of regular input characters, those which comprise the majority of the text to be processed, the main character table, or MCT as it is referred to, is a font table and the entry addressed contains first of all the width value of a character, second, the code for the output character, and finally, flag bits which are useful to the program in processing the character. However, if there is anything out of the ordinary about the character being processed—and this refers to alphanumeric characters as well as machine function characters—the character is classed as an irregular character and the entry in the MCT becomes the address of a special routine designed to deal with that specific irregular character. Typical of irregular characters are the space bar codes, carriage returns, shift codes, and the letters f, l, and i which may have to be combined into ligatures.

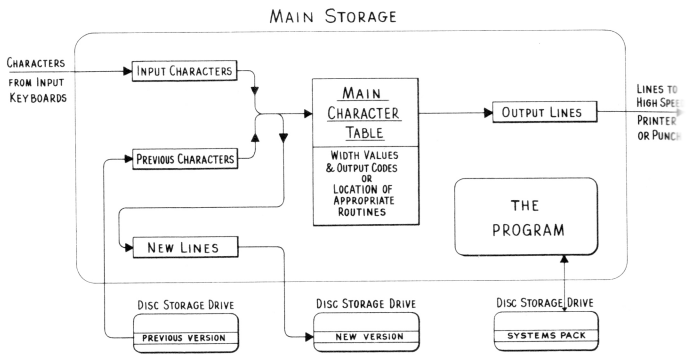

Fig 2 Main storage

The illustration shows two sources of input characters to the MCT. The second source is that of characters obtained from earlier versions of the story being processed. Line numbers entered by the keyboard operator identify the lines which contain the characters to be loaded into the previous characters area. The program then switches to this source and when the requisite number of characters have been processed, the program switches back to the input character area.

As each character is presented to the MCT it is also moved to the new lines area and there it is collected until a line break decision is made. Then these characters are moved to the copy memory file being recorded by one of the disc storage drives. The line being stored is now permanently associated with the line number that is also printed out from the output lines area.

Width values stored in the main character table are relative. They follow the Monotype system of computation since this is virtually identical with the Photon counting system and the absolute widths of Linotype matrices are relatively insensitive to the method of counting as long as a fine enough unit is used. A test is made of the accumulated width in the line for each segment of input characters processed. By redefining 'segment' as the calculations come closer to the end of the line, the program is able to effect decisions as to when to terminate the line. Thus, until the line becomes overset, calculations to test the condition of the line are made only at the spaces. However, the condition of the line at each space is recorded, so if the previous segment calculated was found to be within the justification range this fact is noted. Once having gone over the range the program backs up to the previous segment end. If at this point the line was underset, then obviously the line will not set without hyphenation, and the program breaks away to the hyphenation routine. Then the segment is redefined as the individual syllables of the word whose breaks have been calculated by the hyphenation program. Where hyphenation fails to produce a line which is physically settable, the program will redefine the segment as

being each individual character and will break the line at any character which will set. The tightness of set is determined by the values assigned to the width of space in settable lines. There are four such parameters which can be either held constant or varied from time to time in order to effect the tightness of the set lines.

Controlling the system

One of the most important aspects of any computerized copy processing system is the manner in which the people who work with the system communicate with it. In designing the copy processor every attempt has been made to keep the editorial and keyboard personnel in mind and not to box them by too limited a system of communication. To meet the dual and somewhat contradictory specifications of simplicity and versatility the function code statements which are entered from the keyboard are simplest in form for the most common task while reserving the more complex forms for the relatively infrequent usage. For example, two keystrokes are required to indicate the beginning of a new paragraph—providing the beginning of the paragraph, typographically speaking, is conventional in every respect. Of course, what is conventional for one publication or even a specific department of a given publication may differ from the specification for another department or another publication. However, these differing specifications are contained within the computer's memory, and each take is begun by indicating to the computer the publication being served. Thereafter the same two keystrokes at the start of each paragraph are adequate to produce the correct typeset result. As a next level of variability, if the indents of the entire tape are uniform but different from the standard indent a simple statement at the start of the take will produce the new indents for this one take only. Again the same two-stroke statement is used to produce the new indentation. Finally, if there is to be one or more paragraphs which differ from either the standard or the alternative standard indent, the specific code at the start of each paragraph can be enlarged to specify the exact number of ens for that paragraph only. The same statement can be further enlarged in order to produce hanging indent effects.

With few exceptions function information is transmitted to the computer via the keyboard by means of groups of characters called function code statements. Each statement begins with the phi character—a unique code which has been set aside for this purpose and is obtained by striking a particular key on the teletypesetter keyboard.

One of the most important code statements used by the keyboard operators is the one which enables them to call out solid lines of text from a previous version. The system differs in one respect from other editing systems in that it uses a reconstruction approach in developing the new version. In effect, the keyboard operator retypes the entire new version except that line numbers take the place of full typing where one or more complete lines require no changes. Where such saved lines occur as a block one statement giving the first and last line number of the block is keyboarded. The purpose in using the reconstruction approach is to enable complex reordering of text material to be readily controlled at the keyboard. Thus, if an editor changes the order of two paragraphs, the keyboard operator need only type the line numbers covering the paragraph to appear first in the new version and then to keyboard the line numbers of the paragraph which has been moved down in the text.

A feature that had to be built into the processor to meet the needs of *Time's* editors is a message handling system. A message is loosely defined as any information that must accompany the text, but will not ultimately appear in print. It may be a remark pencilled in the margin of the copy by one editor for the attention of another. It may be a line to the printer specifying or clarifying the typography. It may even be internally generated by the computer to apprise its operators of some unusual condition. But whatever its source, it must not upset the normal line-break calculations. The problem posed is that an operator keying unjustified copy has no way of knowing when to insert a message so that it does not alter the line-breaks. Accordingly, this job is left to the computer. The operator identifies the message by means of a function code statement and all of the message characters are set aside until an appropriate time for their insertion into the output. Messages can appear in any of three different areas of the hard copy or tape depending on the format preferred by the publication. 'In-line' messages are inserted within the same measure and follow directly after the text line they are associated with. They are processed exactly as any other line and so, if left in, the copy will eventually be typeset. 'End-line' messages are similar to in-line except that they are saved out until the end of the take and then set within the same margins. 'Side-line' messages are truly marginal messages. Whatever space remains to the right of the text is filled with message characters and this practice is carried from line to line until the message is fully transcribed.

Thus far we have described the system as it is now being programmed and will be operated starting in 1967. But the appearance of the general purpose computer on the editorial production scene does not represent the culmination of a long series of technical advances so much as it represents the beginning of a completely new era. Typesetting and editing are only the first and probably the most obvious of the operations which will benefit from the use of the computer. While we speak of the copy processor as a system, it is in fact a part of aa much larger system which under control of the editorial worker accepts information in its earliest stages and produces final press-ready master images. We are moving toward the use of composing systems which are not limited to the production of isolated blocks of text but which will be able to produce full pages of composition including half-tones and artwork at rates measured in seconds.

But these 'super composers' of the future are completely wedded to the computer for two reasons. First is the complexity of the makeup job which requires that the signals be presented to the composer in a precise order in order to realize the full value of the composing machine; and further, because these systems are almost invariably based on photo typesetting, they are quite intolerant of errors.